

# Gifted Child Quarterly

<http://gcq.sagepub.com/>

---

## **Myth 16: High-Stakes Tests Are Synonymous With Rigor and Difficulty**

Tonya R. Moon

*Gifted Child Quarterly* 2009 53: 277

DOI: 10.1177/0016986209346945

The online version of this article can be found at:

<http://gcq.sagepub.com/content/53/4/277>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



National Association for Gifted Children

**Additional services and information for *Gifted Child Quarterly* can be found at:**

**Email Alerts:** <http://gcq.sagepub.com/cgi/alerts>

**Subscriptions:** <http://gcq.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://gcq.sagepub.com/content/53/4/277.refs.html>

# Myth 16: High-Stakes Tests Are Synonymous With Rigor and Difficulty

Tonya R. Moon

Curry School of Education, University of Virginia

Sir James Frazer, a social anthropologist in the late 1800s/early 1900s, believed that myths were pre-scientific attempts to interpret the natural world. He suggested that myths provide protoscientific explanations of some phenomenon that has not been fully explored. Merriam-Webster's defines a myth as "a usually traditional story . . . that serves to unfold part of the world view of a people or explain a practice, belief, or natural phenomenon . . . especially one embodying the ideals and institutions of a society of segment of society" (<http://www.merriam-webster.com/dictionary/myth>). The myth equating high-stakes testing with rigor and difficulty, however, is one that can be debunked given the empirical work that has been conducted in this area. To do this requires a basic understanding of the history of high-stakes tests and the evidence to date of their consequential validity.

## History of High-Stakes Tests

High-stakes tests as we know them today originated in the late 1840s in Boston in an attempt to monitor schools' effectiveness by comparing classrooms and schools. During the World War I era, a new emphasis on large-scale testing occurred with the emergence of the large-scale group intelligence test, the Army Alpha Test. During the 1960s the government began monitoring educational systems and student performance with the National Assessment of Educational Progress (NAEP). Alongside NAEP came the 1965 Elementary and Secondary Education Act (ESEA). These two institutions, NAEP and ESEA legislation, were the first to formally use tests for monitoring the performance of the nation's students. Some argue that the use of tests in this way led policy makers and the general public to believe that test scores were reflective of the quality of the public school system.

The 1970s and 1980s saw the enactment of the minimum-competency testing (MCT) era where mandated testing programs ensured that individual students obtained a *minimal level* of proficiency in the area of basic skills. This era had significant negative influences on instruction and marked a belief that instruction can and should be shaped directly by tests. This set the stage for what we now know as the standards movement and the *No Child Left Behind* (NCLB, 2002) legislation.

## Current Educational Testing Context

Although several events and trends (e.g., *A Nation at Risk* [National Commission on Excellence in Education, 1983]; standards movement) during the 1980s and 1990s led to our current accountability system, the major change was the shift from minimum competency to high expectations for *all* students. States redesigned their testing programs shifting from off-the-shelf tests to tests specifically aligned with state content standards. These tests were designed to measure a broader range of skills and specific content knowledge than MCTs, often focusing on skills and knowledge that were easiest to measure through multiple-choice tests rather than critical and higher-order thinking skills and complex aspects of disciplinary content (Webb, 1999). The federal statute, *No Child Left Behind* (NCLB, 2002), pushed the use of tests for accountability purposes, which resulted in a resurgence of high stakes for students and schools. Unintended consequences documented as a result of using assessments for high-stakes accountability purposes include narrowing of the curriculum and instruction, the use of unethical test preparation practices,

**Author's Note:** Please address correspondence to Tonya R. Moon, Curry School of Education, University of Virginia, P.O. Box 400277, Charlottesville, VA 22904-4277; e-mail: [trm2k@virginia.edu](mailto:trm2k@virginia.edu).

inappropriate uses of assessment scores such as reassignment of teachers or principals, and for some students, decreased motivation to learn.

## Empirical Evidence to Debunk the Myth

High-stakes testing as carried out by the NCLB legislation requires states to (among other things):

- Implement content standards for reading/language arts and mathematics for Grades 3 to 8 and high school and corresponding annual assessments by 2005-2006.
- Implement content standards for science in one grade at each of three school levels (Grades 3-5, 6-9; 1-12) by 2005-2006 and corresponding assessments by 2007-2008.

It is well understood in the field of measurement that no single assessment can serve both accountability and classroom instruction goals simultaneously, and do it well (Stiggins, 2003). It has also been acknowledged by the National Research Council (NRC) that there is a need to bridge the gap between large-scale and classroom assessment (NRC, 2003). The inability of high-stakes tests to serve dual purposes (accountability and student learning) results in a phenomenon known as a “default” philosophy of education (Gunzenhauser, 2003), where tests drive the curriculum and limit instructional innovations for the sole purposes of high test scores, this default has dire implications for all students—especially the gifted.

To understand that rigor and difficulty are not taken into consideration with the development high-stakes tests, a working definition of rigor and difficulty should be provided. Principles developed from the Leadership Training Institute (LTI; Passow, 1982) have long been upheld as the enduring principles for appropriately challenging curriculum for the gifted. The basic premise for the seven principles is that the content of curricula should focus on and be organized to include more elaborate, complex, and in-depth study of major ideas, problems, and themes that integrate knowledge within and across systems of thought. This type of curricula allows for the development and application of productive thinking skills enabling students to reconceptualize existing knowledge and/or generate new knowledge (Passow, 1982).

Much research has been done over the past two decades regarding the effects of high-stakes testing on

curriculum. although some research indicates positive effects in terms of focusing the curriculum (Stecher & Chun, 2001), the majority of evidence suggests negative effects in terms of narrowing the curriculum and overemphasizing decontextualized skills (Center on Education Policy, 2006). Specific negative effects include

- adoption of instructional approaches that resemble testing methods and formats (Pedulla et al., 2003); and
- negative coaching where excessive amounts of time are spent on activities that focus on incidental aspects of a test (e.g., Jones et al., 1999). These effects result in diminished exposure to high-quality, rigorous curriculum for all students, including gifted students.

To better understand the disconnect between curriculum focused for preparation on high-stakes tests and the ideals of rigor and difficulty that is appropriately challenging curriculum for gifted children it is helpful to examine a sample released item from one state’s assessment program. For this problem, students simply substitute the value 6 for each  $x$  in the equation to determine the value of the function, a very algorithmic process, and select the correct answer as opposed to the engaging in complex problem solving and reasoning focused on major ideas and problems suggested in the LTI principles (Passow, 1982).

### PROBLEM 1: ALGEBRA I

Read and solve each question. Then mark the space on your answer document for the best answer. For this test you may assume that the value of the denominator is not zero.

If  $f(x) = x^2 + 2x + 3$ , what is the value of  $f(x)$  when  $x = 6$ ?

- A: 27
- B: 42
- C: 51 (correct answer)
- D: 60

## Summary

To completely debunk this myth in gifted education, the field must centralize efforts. We need to consider alternatives to the current system of assessment and the delivery of instruction.

1. Although the focus on high-stakes testing for accountability purposes does not seem to be a passing fad typical of many educational phenomena, it does suggest that leaders in the field and proponents of gifted education need to *become strong voices* for a balanced system that emphasizes assessment for learning with assessment for accountability purposes.
2. School settings must advocate for flexible grouping configurations to accommodate the varying needs of gifted students to master basic skills and concepts as well as opportunities to move beyond the test preparation when evidence warrants.

There is already some enthusiasm being built for this type of system with the Commission on Instructionally Supportive Assessment, a group comprised of five leading school, teacher, and administrator organizations. In a guide for policymakers, the Commission on Instructionally Supportive Assessment (2001) wrote “teachers, principals, and district superintendents will readily embrace accountability measures if they are tied to effective assessment systems designed and implemented to improve classroom instruction” (in the Preface). To move beyond the myth, a balanced approach to assessment must occur, one that includes changes in our actions as well as our beliefs about high-stakes tests.

### Note

See <http://www.maine.gov/education/mea/mearelitems.htm>. These are representative sample items from one state; similar items from other states can be accessed from each state department of education’s Web site.

### References

- Center on Education Policy. (2006). *Ten big effects of the No Child Left Behind Act on Public Schools*. Retrieved January 12, 2008 from <http://www.cepdc.org/document/docWindow.cfm?fuseaction=document.viewDocument&documentid=29&documentFormatId=596>
- Commission on Instructionally Supportive Assessment. (2001, October). *Building tests to support instruction and accountability: A guide for policymakers*. Retrieved August 5, 2009, from <http://www.testaccountability.org/>
- Gunzenhauser, M. G. (2003). High-stakes testing and the default philosophy of education. *Theory Into Practice*, 42, 51-58.
- Jones, G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81, 199-203.
- National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: Author.
- National Research Council. (2003). *Assessment in support of instruction and learning: Bridging the gap between large-scale and classroom assessment*. (Workshop report, Committee on Assessment in Support of Learning). Washington, DC: National Academies Press.
- No Child Left Behind Act of 2002, Pub Law no. 107-110 (2002, January). Retrieved January 12, 2009 from <http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>.
- Passow, A. H. (1982). Differentiated curricula for the gifted/talented. In *Curricula for the gifted: Selected proceedings for the First National Conference on Curricula for the Gifted/Talented* (pp. 4-20). Ventura, CA: National/State Leadership Training Institute on the Gifted and Talented.
- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived effects of state-mandated programs on teaching and learning: Findings from a national survey of teachers*. Boston: National Board on Educational Testing and Public Policy.
- Stecher, B. M., & Chun, T. (2001). *School and classroom practices during two years of education reform in Washington state* (CSE Tech. Rep. No. 550). Los Angeles: University of California, Los Angeles, Center for Research on Evaluation, Standards, and Student Testing.
- Stiggins, R. (2003). *Balanced assessment: The key to accountability and improved student learning*. Washington, DC: National Education Association.
- Webb, N. L. (1999). *Alignment of science and mathematics standards and assessment in four states*. (Research Monograph No. 18). Arlington, VA: National Science Foundation.

**Tonya R. Moon**, PhD, is an associate professor of education at the University of Virginia where she teaches graduate courses in the area of research design. Her current research focuses on increasing elementary teachers’ content knowledge in the areas of science, technology, and mathematics as well as increasing student engagement and achievement in these areas through Project Parallax. She is also a principal investigator for the National Research Center on the Gifted and Talented where she is collaborating on an elementary reading study investigating a new reading curriculum. She also chairs of the University of Virginia’s Institutional Review Board for the Behavioral and Social Sciences. In 2003 she was given the Early Scholar Award by the National Association for Gifted Children.